ANALYZING THE EFFECTS OF UNFORCED NATURAL VARIABILITY AND ANTHROPOGENIC FORCING ON ENSO VARIABILITY USING CESM

BY

BENJAMIN AARON VEGA-WESTHOFF

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Atmospheric Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Adviser:

Assistant Professor Ryan Sriver

ABSTRACT

The El Niño-Southern Oscillation (ENSO) is an important contributor to Earth's inter-annual climate variability, with worldwide weather effects (*Whetton and Rutherfurd*, 1994; *Hoerling and Zhong*, 1997; *Dai and Wigley*, 2000). Understanding how ENSO may change with climate is a major challenge, given the internal variability of the system and relatively short observational record (*Wittenberg*, 2009). Much recent research has used multi-model ensembles to address the effects of climate change on ENSO (*Stevenson*, 2012; *Cai et al*., 2014; *Kim et al*., 2014). Here we analyze ENSO in a Community Earth System Model (CESM) ensemble that samples internal variability of the coupled Earth system. We present results from a 50-member climate change ensemble experiment, using historical climate forcings (1850-2005) and projections to 2100 following the representative concentration pathway 8.5 (RCP8.5). With this ensemble, and a ~5000 year control run with constant pre-industrial conditions, we examine ENSO variability under different forcing regimes. We compare the effects of anthropogenic climate change with the effects of natural modulations on ENSO sea surface temperature (SST). We find that any changes in ENSO SST due to climate change are secondary to natural modulations.

ACKNOWLEDGEMENTS

I would like to thank my adviser Professor Ryan Sriver for his support and guidance on this project. He also provided the CESM output used in the study. The CESM project is supported by the National Science Foundation and the Office of Science (BER) of the U.S. Department of Energy. ERSST v3b data is provided by NOAA and hosted by NCEI (https://www.ncdc.noaa.gov/data-access/). The World Climate Research Programme's Working Group on Coupled Modeling is responsible for CMIP and the U.S. Department of Energy's program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. CMIP5 data was obtained from the KNMI Climate Explorer (https://climexp.knmi.nl/). I thank group members and officemates for their advice and feedback, Gan Zhang for helpful comments, Jake Mulholland for proofreading, and finally my wife for her support.

TABLE OF CONTENTS

1. INTRODUCTION.....1

2. MODEL RESULTS AND DISCUSSION.....3

3. METHODS.....8

4. SUPPLEMENTARY ANALYSES.....11

REFERENCES.....16

FIGURES.....18

TABLES.....28

1. INTRODUCTION

Natural variations within Earth's climate system are an important source of uncertainty in projections of future climate, particularly on interannual to decadal timescales (*Hawkins and Sutton*, 2009). They can influence variability in key climate change indicators, such as regional temperature and precipitation patterns (*Hawkins and Sutton*, 2011), and global temperature trends (*Huber and Knutti*, 2014). Earth system modeling approaches that capture the effect of natural variability, or more precisely the variability within the coupled system in the absence of time-varying external forcing, can be used to characterize the importance of internal variability when considering potential anthropogenic climate trends (*Deser et al*., 2012). This type of approach is especially useful for ENSO, which undergoes decadal modulations that have hindered even qualitative predictions of amplitude change in the coming decades.

ENSO can be characterized by the Niño3.4 index, the climatological SST anomaly in the Niño3.4 region of the equatorial Pacific (outlined in Fig. 1). El Niño (La Niña) events correspond to a positive (negative) Niño3.4 index, with the strength of the event tied to the magnitude of the anomaly. Several ENSO metrics can be derived from the Niño3.4 index time series: the spectrum reveals the frequency distribution of events, skewness roughly measures the relative strength of El Niño and La Niña events, and the standard deviation is tied to ENSO SST amplitude.

In climate model simulations with constant pre-industrial forcing, ENSO frequency and SST amplitude undergo strong decadal modulations. Several hundred years of data are required to sample these modulations under a given forcing (*Wittenberg*, 2009; *Stevenson et al*., 2010). In the most recent multi-model ensemble, there is no consensus change in ENSO under global warming (*Stevenson*, 2012, *Cai et al*., 2014), and if there are robust changes, they may be time-dependent (*Kim et al*., 2014). However, the ensemble does display robust changes in other key ENSO properties, such as the associated precipitation anomalies and event propagation characteristics (*Power et al*., 2013, *Santoso et al*., 2013, *Cai et al*., 2014).

2. MODEL RESULTS AND DISCUSSION

2.1 Ensemble and Model description

Here, rather than a multi-model approach that samples model uncertainty in the processes that govern ENSO, we analyze a large ensemble from a single model. The ensemble samples only this model's internal variability, which we define as its year-to-year variations in the absence of time-varying greenhouse gas forcing. Our ensemble uses a low-resolution configuration of CESM (T31x3, described in the methods section). Individual ensemble members branch off every 100 years from a fully-coupled, equilibrated, control simulation. Each member is then forced with historical values through 2005, then RCP8.5 until 2100. Because each member is identically forced after starting at a unique time in the equilibrated control, any differences between members are solely due to internal variability of the coupled climate system.

While low-resolution CESM exhibits mean state biases in equatorial Pacific SST, it does capture realistic variability associated with ENSO (*Shields et al*., 2012, our Fig. 1). In this regard it is comparable with members of the latest multi-model ensemble (*Jha et al*., 2014). Low-resolution CESM reproduces tropical Pacific mean state seasonality, as well as ENSO's seasonal phase locking (Figs. 2, 3). Low-resolution CESM also simulates the recharge-discharge mechanism of the equatorial Pacific, in which changes in ocean heat content accumulate before ENSO events occur (Fig. 4; *Jin*, 1997; *McPhaden et al*., 2006).

2.2 Results

We examine ENSO in three different forcing regimes: constant pre-industrial forcing, 1940-1990 historical forcing, and RCP8.5 2040-2090 projected forcing. Each forcing regime ensemble has 50 independent members sampling ENSO's natural modulations. Figure 5 shows the ENSO spectra for each forcing regime. The median spectrum for 1940-1990 roughly matches the corresponding observed spectrum, though the power at longer periods is underestimated. Each ensemble exhibits significant spread in spectral shape, with some members peaking around four years and others at less than two years, while the normalized power in those peaks can vary by greater than a factor of four. The medians of each forcing regime ensemble, on the other hand, are very similar in shape, and the 95% confidence intervals of the ensemble means overlap over much of the range (Fig. 6). Differences between individual members, due to ENSO's natural modulations, are much larger than the differences between the ensembles themselves, which would be due to their different forcing regimes.

Figure 7 compares the distributions of several statistical properties of the Niño3.4 index in the different forcing regime ensembles. Standard deviation, an indicator of ENSO SST amplitude, is somewhat underestimated in the model. The observed 1940-1990 value is above the 95th percentile of the 1940-1990 ensemble. The distributions for each forcing regime have similar widths and sizeable overlap. The distributions of the latest inter-model

ensemble are also shown. These distributions, sampling a mix of internal variability and model uncertainty, are wider than those of the CESM ensemble (Fig. 7a).

Skewness roughly indicates the relative strength of El Niño and La Niña events. The observed 1940-1990 skewness is positive (El Niño events are generally stronger than La Niña), and closely matches the CESM ensemble median over the same time period. Similar to standard deviation, there is sizeable overlap between the CESM forcing regime ensembles (Fig. 7b). The CESM ensemble distributions are again comparable for El Niño and La Niña event counts (as defined in the Methods section) (Fig. 7c, d).

Given approximate normality (see Q-Q plots in Fig. 8), we perform two-sided t-tests comparing the pre-industrial ensemble mean with that of the 2040-2090 ensemble mean for each statistical property. None of the differences are significant at 95% confidence, even given the good detection sensitivity allowed by this large dataset (e.g., we could have rejected the hypothesis of equal means for a change in standard deviation greater than ~4%) (Table 1).

For an explicit analysis of ENSO evolution, we next examine the continuous record of 20-year running ENSO statistical properties. Individual simulations under both pre-industrial and climate change forcings exhibit ENSO SST amplitude (i.e., Niño3.4 index standard deviation) modulations comparable with those in the observations. Modulations of ENSO skewness larger than those in the observations occur under both pre-industrial and climate change forcings. Over the combined historical and projection time period, any ensemble change in either statistical property is negligible compared with natural modulations (Fig.

9). We repeat this analysis for the Niño3 and Niño4 regions, which are sensible to different flavors of ENSO and may therefore be differently affected by climate change (*Kim and Yu*, 2012)(Figs. 10, 11). These regions also exhibit natural modulations larger than any ensemble changes. However, unlike Niño3.4 and Niño3, the ensemble Niño4 amplitude is significantly lower in 2040-2090 compared to pre-industrial (2040-2090 is 5% lower, p-value = 0.002).

2.3 Discussion

When the goal is realistic projection of ENSO evolution, single-model ensembles are certainly not as valuable as those from a large multi-model intercomparison. Multi-model ensembles sample parameter and structural uncertainties that reflect our uncertainty about the Earth system and are thus less susceptible to model biases. On the other hand, without those confounding uncertainties, a single-model ensemble permits straightforward characterization of ENSO decadal modulation and comparison of this modulation with possible changes due to climate change.

Moreover, a large ensemble allows characterization through time of ENSO's decadal modulations, which are not *a priori* climate change independent. A change in these modulations could have important implications for ENSO impacts (e.g., compare two futures with the same time-averaged ENSO SST amplitude: one in which ENSO becomes completely regular and one in which decadal modulation doubles).

Several hundred years are needed for reasonable sampling of simulated ENSO modulations under constant forcing. These modulations must be accounted for when analyzing ENSO under projected climate change. A change in ENSO in a single simulation may only reflect natural modulations, rather than a robust response of the model to climate change (e.g., Fig. 12). In a large enough ensemble, sampling the decadal variability of the coupled climate system, these modulations can be averaged out.

3. Methods

3.1 CESM Experiment

The experiment uses the fully-coupled low-resolution configuration of the Community Earth System Model (CESM), T31x3. The atmospheric model component has a spectral resolution of ~3.75° x 3.75° and 26 vertical levels. The ocean model component has a nominal horizontal resolution of 3° (changing to less than 1° near the equator) and 60 vertical levels. While the low-resolution version of CESM has several climate biases, it captures tropical Pacific interannual temperature variability associated with ENSO and other properties related to ENSO (*Shields et al*., 2012, our Figs. 1, 2, 3, 4).

The experiment first undergoes a ~4000 year equilibration phase with pre-industrial forcings, allowing the deep ocean to reach near-dynamic equilibrium. After equilibration, the control simulation continues for another 5000 years. 50 simulations branch off at 100-year intervals of this equilibrated control. These 50 simulations run from 1850-2100, using historical anthropogenic and natural forcings between 1850 and 2005 and RCP8.5 from 2006 to 2100 (*Moss et al*., 2010). The 50-member ensemble, with identical forcings, only samples unique initialization times in the equilibrated control (*Sriver et al*., 2015). Thus, differences between members reflect uncertainty due to joint internal unforced variability of the fully-coupled system (ocean, atmosphere, land, and sea-ice components).

We further break up the ensemble into a 50-member 1940-1990 ensemble and a 50-member 2040-2090 ensemble. Independent sections of the equilibrated, control simulation make up a 50-member control ensemble. Thus, we have three ensembles that sample the internal variability in three different forcing regimes. Repeating the analysis with longer time sections did not change our results (not shown).

3.2 Statistical Analyses

We focus on SST in the Niño3.4 region, (5°S–5°N, 170°–120°W). We calculate a Niño3.4 index as the monthly average SST climatological anomaly in the region. To isolate ENSO variability from long-term warming, a 211-month triangle smoothing is subtracted from the index time series (as in Fig. 1 of *Wittenberg*, 2009). The resulting trend-removed Niño3.4 index time series has no warming trend.

For spectral analysis, we calculate the maximum entropy power spectrum (*Burg*, 1967) for each trend-removed time series. In order to visually isolate the ENSO peak, the spectrum is normalized by frequency.

For box-and-whisker plots, the notches extend to $\pm(1.58 * IQR)/\sqrt{n}$, where IQR is the interquartile range and n is the number of ensemble members (50 for CESM, 35 for CMIP5). The notches give a rough 95% confidence interval for the difference between two medians (*Chambers et al.*, 1983). El Niño (La Niña) events are identified by at least 5 consecutive

months in which the three-month average of the trend-removed index is greater (less) than 0.5°C (-0.5°C), similar to the method used by NOAA Climate Prediction Center.

4. Supplementary Analyses

4.1 Further verification of low-resolution CESM ENSO

It is reassuring that the amplitude and periodicity of ENSO in low-resolution CESM are comparable with observations (*Shields et al*., 2012, Figs. 5, 7). However, ENSO is dependent on a series of coupled oceanic and atmospheric feedbacks. Error compensation among these feedbacks can lead to an apparently realistic ENSO and projection overconfidence (*Bellenger et al*., 2014, *Kim et al*., 2014). Here we present several additional verifications of the modeled ENSO.

The tropical Pacific mean state, affected by the same feedbacks governing ENSO, can be used as an indicator of a model's ability to capture ENSO (*Bellenger et al*., 2014). Low-resolution CESM captures the zonal asymmetry of tropical Pacific seasonality (Fig. 2).

ENSO variability is observed to peak from November-January. About half of all CMIP3 and CMIP5 models share this peak. Those that capture this phase-locking more accurately capture the temperature-dependence of Eastern Pacific convective regimes (*Bellenger et al.*, 2014), which is an important ENSO feedback (*Lloyd et al*., 2012). Low-resolution CESM's ENSO shares the observed phase locking (Fig. 3).

ENSO can be partially understood as a recharge-discharge of equatorial ocean heat content (*Jin*, 1997). After a build-up of ocean heat, an El Niño event then releases that heat. The 20°C isotherm depth can be used as an indicator of ocean heat content. The cross-correlation of Eastern equatorial Pacific SST and equatorial Pacific 20°C isotherm depth allows insight into whether a model can capture the recharge-discharge ENSO mechanism (*McPhaden et al*., 2006). Observational analysis shows a positive cross-correlation value at negative lags, which means that equatorial ocean heat increases lead increases in Eastern Pacific SST (*Meinen and McPhaden*, 2000). Low-resolution CESM reproduces this peak at negative lags, indicating some recharge-discharge mechanism governing ENSO (Fig. 4).

4.2 Mean State Changes and their Effect on ENSO

Recent analyses of CMIP5 have shown that, while changes in frequency and anomalous temperatures associated with ENSO are uncertain, the precipitation anomalies associated with El Niño and La Niña may become more extreme in the future (*Cai et al.,* 2014, *Cai et al*., 2015). These more frequent extreme events are linked to robust changes in the mean state of the CMIP5 ensemble equatorial Pacific. Under RCP8.5, the CMIP5 ensemble undergoes preferential warming of the Eastern equatorial Pacific. The lower east-west background temperature gradient means that convection in the Eastern equatorial Pacific is more frequently established during El Niño events (*Cai et al*., 2014). The low-resolution CESM ensemble also undergoes preferential warming of the Eastern Pacific under RCP8.5 (Fig. 13a).

Additionally, this preferential warming is correlated with future changes in ENSO variability: the stronger the preferential warming of the East, the more likely ENSO variability increases (Fig. 13b). This correlation is consistent with previous studies, which have found that a more active ENSO is associated with warmer east Pacific temperatures (e.g., *Ogata et al.,* 2013; *Wittenberg*, 2015).

4.3 ENSO Flavor Changes

There is recognition that there are two 'flavors' of ENSO: the Eastern-Pacific type, in which SST anomalies are centered in the East, and the Central-Pacific type, in which SST anomalies occur near the International Date Line (e.g. *Ashok et al*., 2007; *Yu and Kao*, 2007). These two flavors can have distinct teleconnections (*Ashok et al*., 2007). Their underlying dynamics may also be different (*Kao and Yu*, 2009). Therefore, they may be affected differently by climate change (*Kim and Yu*, 2012). We perform a simple check of whether the two flavors are differently affected by climate change in the CESM ensemble. Our original analysis uses SST anomalies in the Niño3.4 region (5°S–5°N, 170°–120°W) as an ENSO index. We repeat the analysis using SST anomalies in the Niño3 (5°S-5°N, 150°W-90°W) and Niño4 (5°S-5°N, 160E°W-150°W) regions. The Niño3 region is affected more strongly by Eastern-Pacific events, while Niño4 is more strongly affected by Central-Pacific events. Differences in climate change effects on the two regions may indicate differences in climate change effects on the two ENSO flavors. As with the Niño3.4 index, internal variability dominates changes due to anthropogenic forcing in both regions (Figs. 10, 11).

However, averaging over all 50 members, there is a significant decrease in Niño4 variability from control to 2040-2090 (two-sided t-test, p-value = 0.0022).

4.4 Additional Spectral Analysis

In the main text, we compare the ensemble ENSO spectra for three forcing regimes/time periods: the unforced/control ensemble, a historical/1940-1990 ensemble, and a projections/2040-2090 ensemble. This comparison is done with a three-panel plot of the spectra and the 5^{th}, 50^{th}, and 95^{th} percentiles of each distribution. Here, we add a plot of the 95% confidence interval for the mean spectrum of each time period ensemble, assuming a t distribution at every frequency (Fig. 6). There is sizeable overlap of the confidence intervals at most frequencies.

4.5 Additional statistical comparison

In the main text, we compare the ENSO standard deviation, skewness, and count of El Niño events and La Niña events in our three time periods/forcing regimes using box-whisker plots. Here we perform a t-test of the hypothesis that the control and 2040-2090 ensembles have equal means (Table 1). Before using this t-test, we first verify that the statistical properties have normal distributions using quantile-quantile plots (Fig. 8). After confirming approximate normality, we use Welch's two-sample t-test. We find that none of

the statistical properties examined has a two-sided p-value less than 0.05, that is, we cannot reject the null hypothesis of equal means with 95% certainty. Comparing control with 1940-1990 and 1940-1990 with 2040-2090 (not shown in table), there is one p-value less than 0.05: the 1940-1990 vs. 2040-2090 La Niña counts (p-value = 0.044). However, given the 12 t-tests performed, a single p-value this small would not be unlikely when sampling from identical normal distributions. Given this fact and the much larger p-values for other La Niña count comparisons, we don't focus on this result.

In Table 1 we also estimate the difference in the means that would be required to reject the null hypothesis with 95% certainty. This detection sensitivity is calculated assuming the ensemble spread of the given property is constant and that there are 98 degrees of freedom. As an example, the 95% detection sensitivity of the ENSO standard deviation is 3.7%, so we would reject the null hypothesis of equal means with 95% certainty if the change in standard deviation were larger than 3.7%.

4.6 Importance of the Ensemble Approach

Given ENSO's considerable natural modulations, it can be dangerous to use a single simulation as indication of ENSO's response to climate change. As a demonstration, we present two CESM ensemble members with particularly divergent ENSO amplitude trends in the 21st century (Fig. 12).

REFERENCES

Ashok, K., Behera, S., Rao, A. S., Weng H., & Yamagata, T. El Niño Modoki and its teleconnection. *J. Geophys. Res.* **112,** C11007 (2007).

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., & Vialard, J. ENSO representation in climate models: from CMIP3 to CMIP5. *Clim. Dynam.* **42,** 1999-2018 (2014).

Burg, J. P. Maximum entropy spectral analysis. Reprinted in *Modern Spectrum Analysis*, D. G. Childers (ed.), 42–48 (IEEE Press, 1978).

Cai W. *et al*. Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Clim. Change* **4,** 111-116 (2014).

Cai, W. *et al.* Increased frequency of extreme La Niña events under greenhouse warming. *Nat. Clim. Change* **5,** 132-137 (2015).

Chambers, J. M., Cleveland, W. S., Kleiner, B. & Tukey, P. A. *Graphical Methods for Data Analysis* (Wadsworth & Brooks/Cole, 1983).

Dai, A. G. & Wigley, T. M. L. Global Patterns of ENSO-Induced Precipitation. *Geophys. Res. Lett.* **27,** 1283-1286 (2000).

Deser, C., Phillips, A., Bourdette, V. et al. Uncertainty in climate change projections: the role of internal variability. *Clim. Dyn.* **38**, 527-546 (2012).

Hawkins E. & Sutton, R. The potential to narrow uncertainty in projections of regional precipitation change. *Clim. Dyn.* **37**, 407-418 (2011).

Hawkins, E. & Sutton, R. The potential to narrow uncertainty in regional climate predictions. *Bull. Am. Meteorol. Soc.* **90,** 1095–1107 (2009).

Hoerling, M., Kumar, A., and Zhong, M. El Niño, La Niña, and the nonlinearity of their teleconnections. *J. Clim.* **10,** 1769–1786 (1997).

Jha, B., Hu, Z.-Z., & Kumar, A. SST and ENSO variability and change simulated in historical experiments of CMIP5 models. *Clim. Dyn.* **42,** 2113-2124 (2014).

Jin, F.-F. An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *J. Atmos. Sci.* **54,** 811–829 (1997).

Kao, H.-Y., & Yu J.-Y. Contrasting eastern-Pacific and central-Pacific types of El Niño, *J. Clim.* **22,** 615–632 (2009).

Kim, S. T. & Yu, J. Y. The two types of ENSO in CMIP5 models. *Geophys. Res. Lett.* **39.11,** (2012).

Kim, S. T., *et al.* Response of El Niño sea surface temperature variability to greenhouse warming. *Nat. Clim. Change* **4,** 786–790 (2014).

Lloyd, J., Guilyardi, E., & Weller, H. The role of atmosphere feedbacks during ENSO in the CMIP3 models, Part III: the shortwave feedback. *J. Clim.* **25,** 4275–4293 (2012).

McPhaden, M. J., Zebiak S. E. & Glantz M. H. ENSO as an integrating concept in Earth science. *Science* **314,** 1740–1745 (2006).

Meinen, C. S., & McPhaden, M.J. Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña. *J. Clim.* **13,** 3551–3559 (2000).

Moss, R. H. *et al*. The next generation of scenarios for climate change research and assessment. *Nature* 463, 747–756 (2010).

Ogata , T., S.-P. Xie, A. Wittenberg, and D.-Z. Sun, Interdecadal amplitude modulation of El Niño/Southern Oscillation and its impacts on tropical Pacific decadal variability. *J. Clim.* **26**, 7280-7297 (2013).

Power, S., *et al.* Robust twenty-first-century projections of El Niño and related precipitation variability. *Nature* **502.** 541-545 (2013).

Santoso, A. *et al*. Late-twentieth-century emergence of the El Niño propagation asymmetry and future projections. *Nature* **504,** 126-130 (2013).

Shields, C. A. *et al.* The low-resolution CCSM4. *J. Clim.* **25,** 3993–4014 (2012).

Smith, T. M., Reynolds, R. W., Peterson, T. C., & Lawrimore, J. Improvements to NOAA's historical merged land–ocean temperature analysis (1880–2006). *J. Clim.* **21,** 2283–2296 (2008).

Sriver, R. L., Forest, C. E. & Keller, K. Effects of initial conditions uncertainty on regional climate variability: An analysis using a low-resolution CESM ensemble. *Geophys. Res. Lett.* **42,** 5468-5476 (2015).

Stevenson, S. L. Significant changes to ENSO strength and impacts in the twenty-first century: Results from CMIP5. *Geophys. Res. Lett*. **39,** L17703 (2012).

Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B., Yeager, S.G. ENSO model validation using wavelet probability analysis. *J. Clim.* **23,** 5540-5547 (2010).

Whetton, P. & Rutherfurd, I. Historical ENSO teleconnections in the eastern hemisphere. *Clim. Change* **28,** 221-253 (1994).

Wittenberg, A. T. Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.* **36,** L12702 (2009).

Wittenberg, A. T. Low-frequency variations of ENSO. *U.S. CLIVAR Variations* **13** (1), 26-31 (2015).

Yu, J.-Y., & Kao, H.-Y. Decadal changes of ENSO persistence barrier in SST and ocean heat content indices: 1958–2001, *J. Geophys. Res.*, **112,** D13106 (2007).

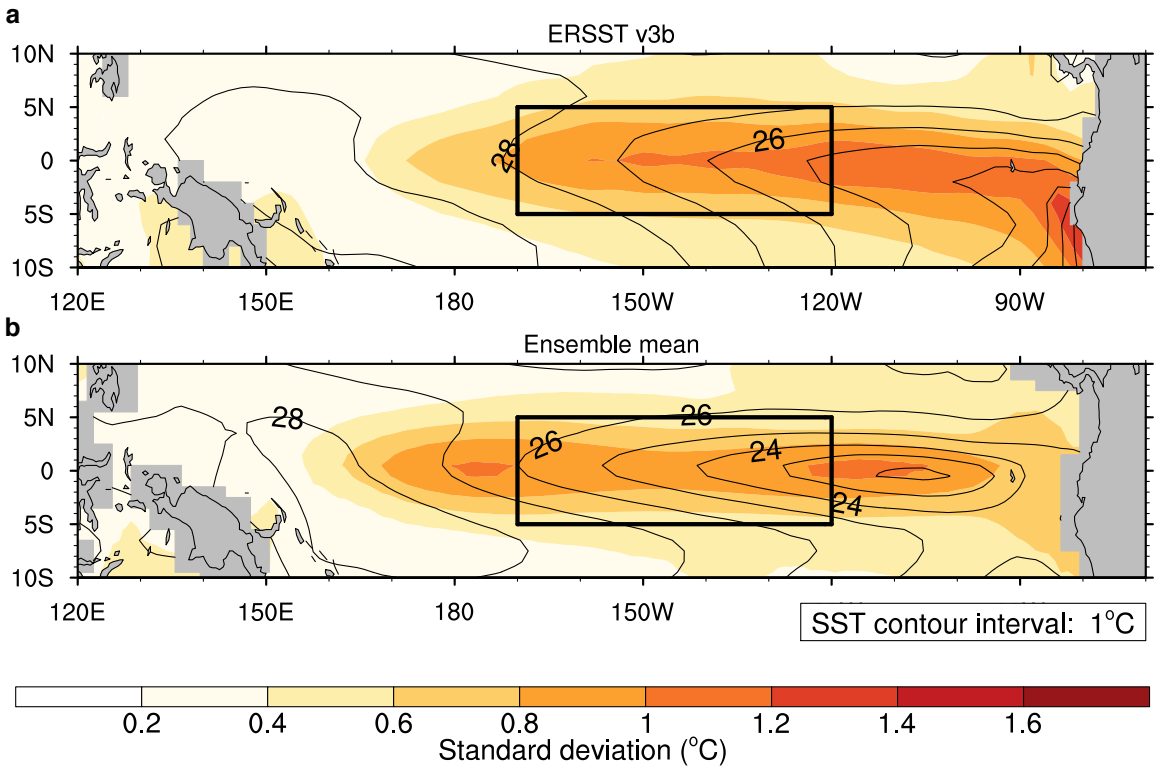

Figure 1 | Tropical mean state and SST variability in observations and CESM. Contours of mean SST (lined) and standard deviation of SST climatological anomalies (colored) for **a**, reconstructed observations from 1940-1990 (ERSST v3b, *Smith et al.*, 2008**)**; **b**, the mean of the 50 CESM ensemble members from 1940-1990. The Niño3.4 region is outlined in black.

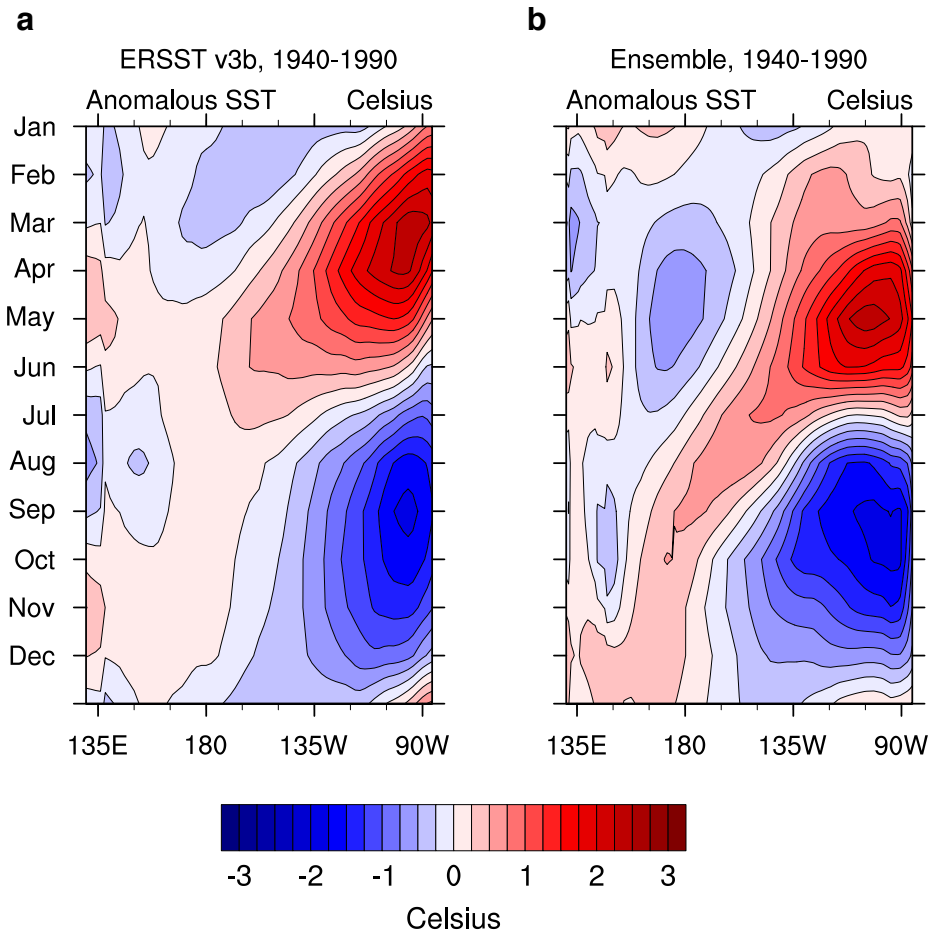

Figure 2 | Annual cycle of equatorial Pacific surface temperature in observations and the CESM ensemble. Longitude-time plots of the annual cycle of surface temperature anomalies in the equatorial Pacific (5°S–5°N) relative to the annual meridional mean from 1940-1990 for **a**, reconstructed observations (ERSST v3b) and **b**, the CESM ensemble average.

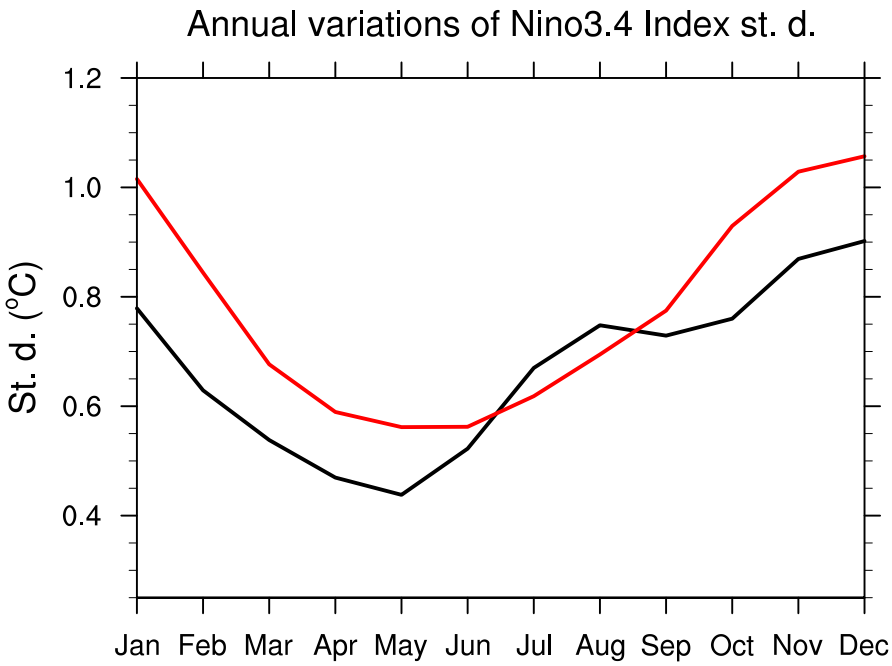

Figure 3 | ENSO phase-locking in observations and the CESM ensemble. Monthly average standard deviation of trend-removed Niño3.4 index from 1940-1990 for reconstructed observations (red), and the CESM ensemble average (black)**.**

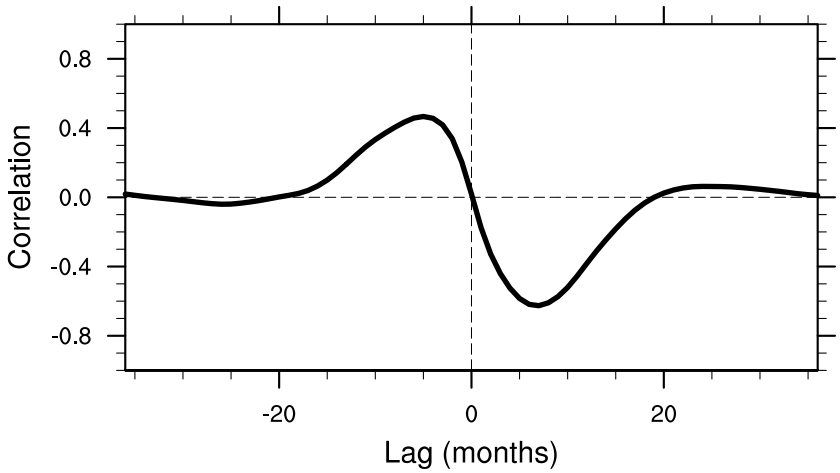

Figure 4 | Lagged cross-correlation of trend-removed Niño3 (5°S-5°N, 150°W-90°W) index and 20°C isotherm depth averaged over the equatorial Pacific (5°S-5°N, 120°E-90°W). Averaged over 1940-1990 of the CESM ensemble mean. Negative lags indicate that 20° isotherm depth anomalies lead Niño-3 SST anomalies.

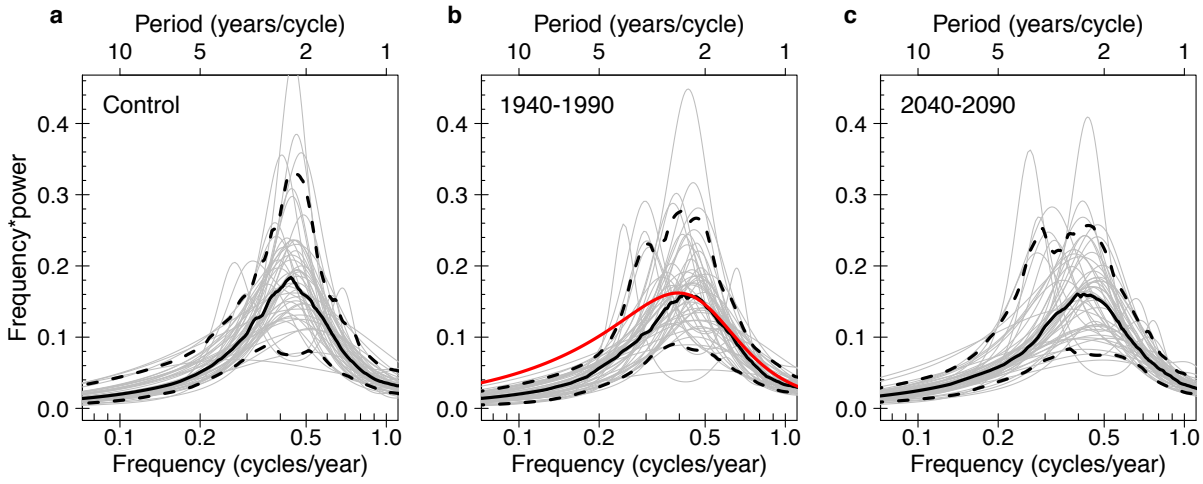

Figure 5 | Maximum entropy power spectra of the trend-removed Niño3.4 index under different forcing regimes. a, Fifty 50-year sections of the unforced, control CESM simulation; **b**, the 50 CESM ensemble members (1940-1990) and, in red, the 1940-1990 reconstructed observational data (ERSST v3b); **c**, the 50 CESM ensemble members (2040-2090). Individual members are shown in grey, dashed curves are the 5% and 95% percentiles, and the solid black curve is the median.

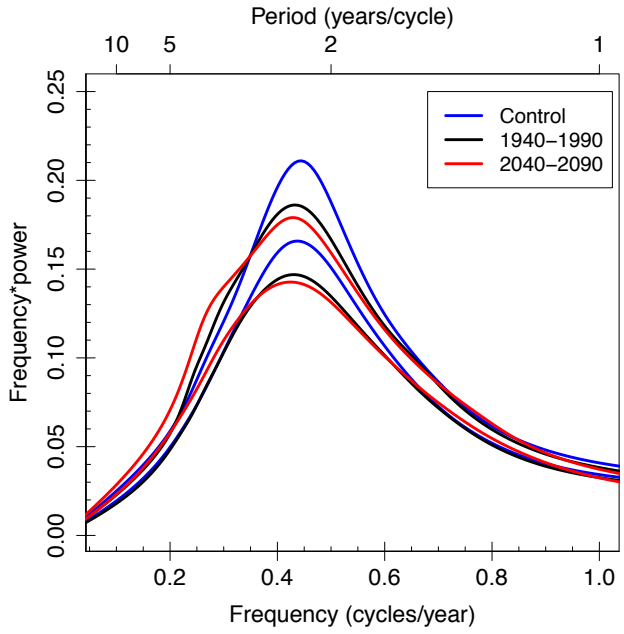

Figure 6 | 95% confidence intervals for the mean maximum entropy power spectrum of the trend-removed Niño3.4 index under different forcing regimes. Fifty 50-year sections of the unforced, control CESM simulation (blue), the 50 CESM ensemble members from 1940-1990 (black), the 50 CESM ensemble members from 2040-2090 (red).

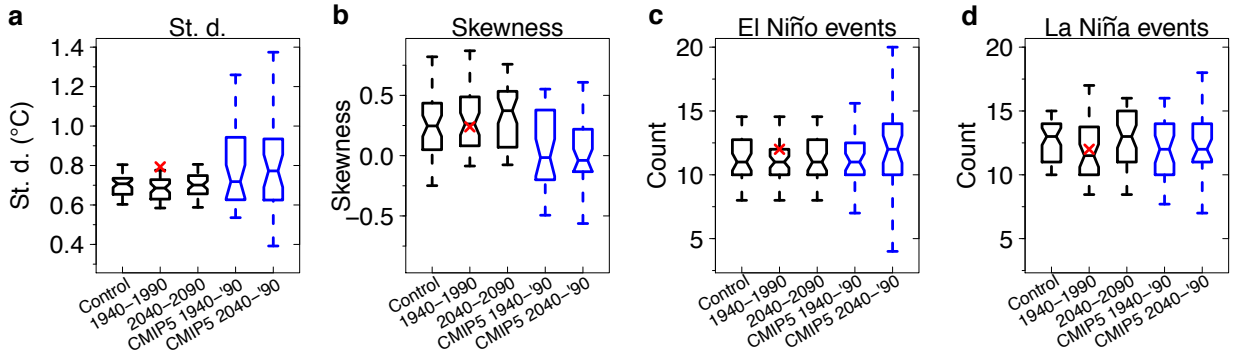

Figure 7 | Box-whisker plots of ENSO statistics under different forcing regimes. a, Trend-removed Niño3.4 index standard deviation; **b**, skewness; **c**, El Niño event counts; **d**, La Niña event counts. The plots show 5th, 25th, 50th, 75th, and 95th percentiles for fifty 50-year sections of the unforced, control CESM run (Control), the 50 CESM ensemble members from 1940-1990 (1940-1990), and the 50 CESM ensemble members from 2040-2090 (2040-2090). In blue are the distributions for the 35-member CMIP5 ensemble. The red x denotes the 1940-1990 reconstructed observational value (ERSST v3b). The notches indicate a rough 95% confidence interval for the difference between two medians.

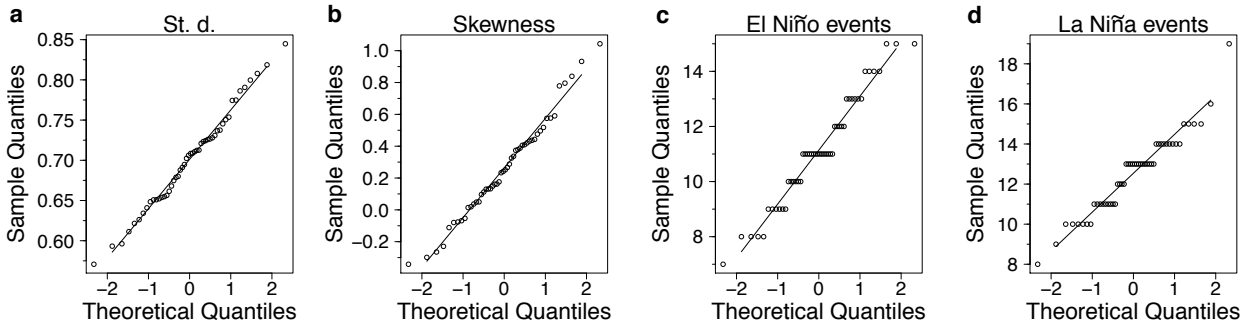

Figure 8 | Quantile-quantile plots comparing control ensemble distributions on the y-axis with normal distributions on the x-axis. a, Standard deviation; **b**, skewness; **c**, El Niño event counts; **d**, La Niña event counts. If a distribution were perfectly normal, its quantile-quantile plot would be perfectly linear.

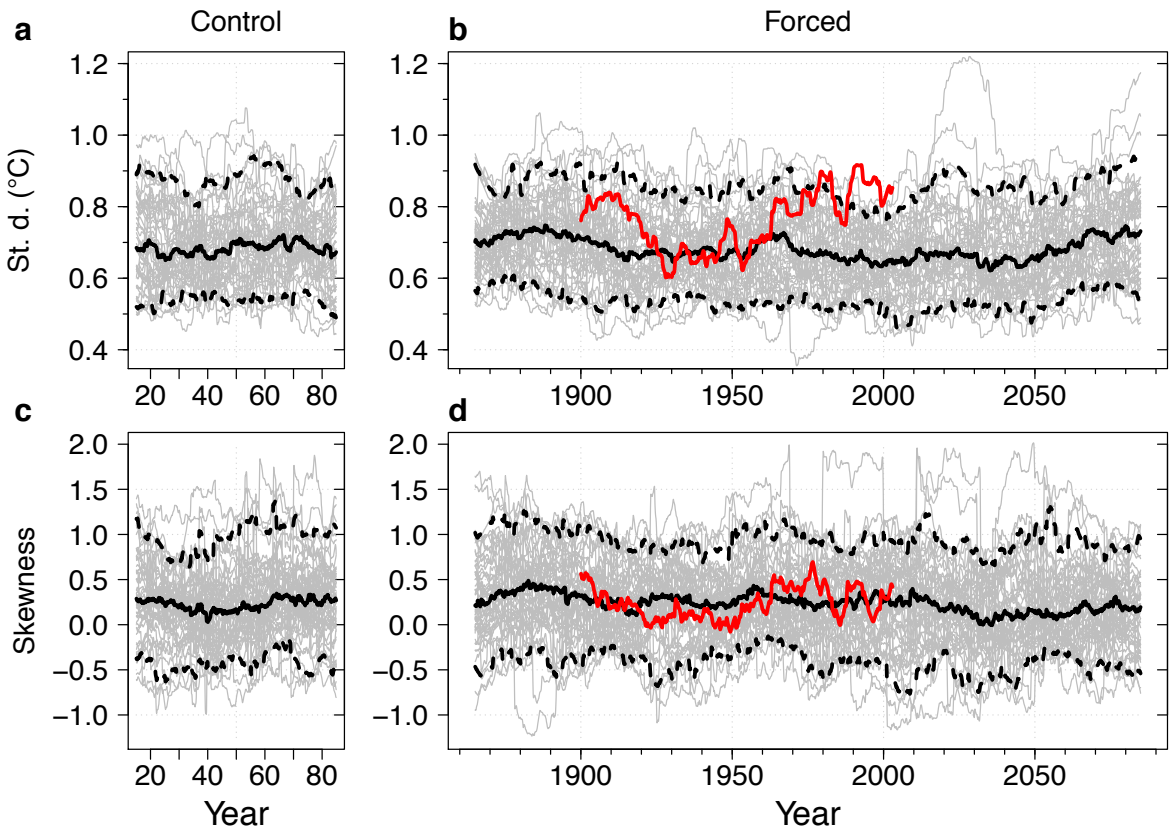

Figure 9 | Running ENSO statistics in the CESM ensemble. a, **b**, 20-year running standard deviation of the trend-removed Niño3.4 index; **c**, **d**, skewness. **a**, **c**, 50 sections of the unforced, control CESM run; **b**, **d**, the 50 CESM ensemble members. Individual members are shown in grey, dashed lines are the 5th and 95th percentiles, the solid black line is the median, and the solid red line is the reconstructed observational record (ERSST v3b).

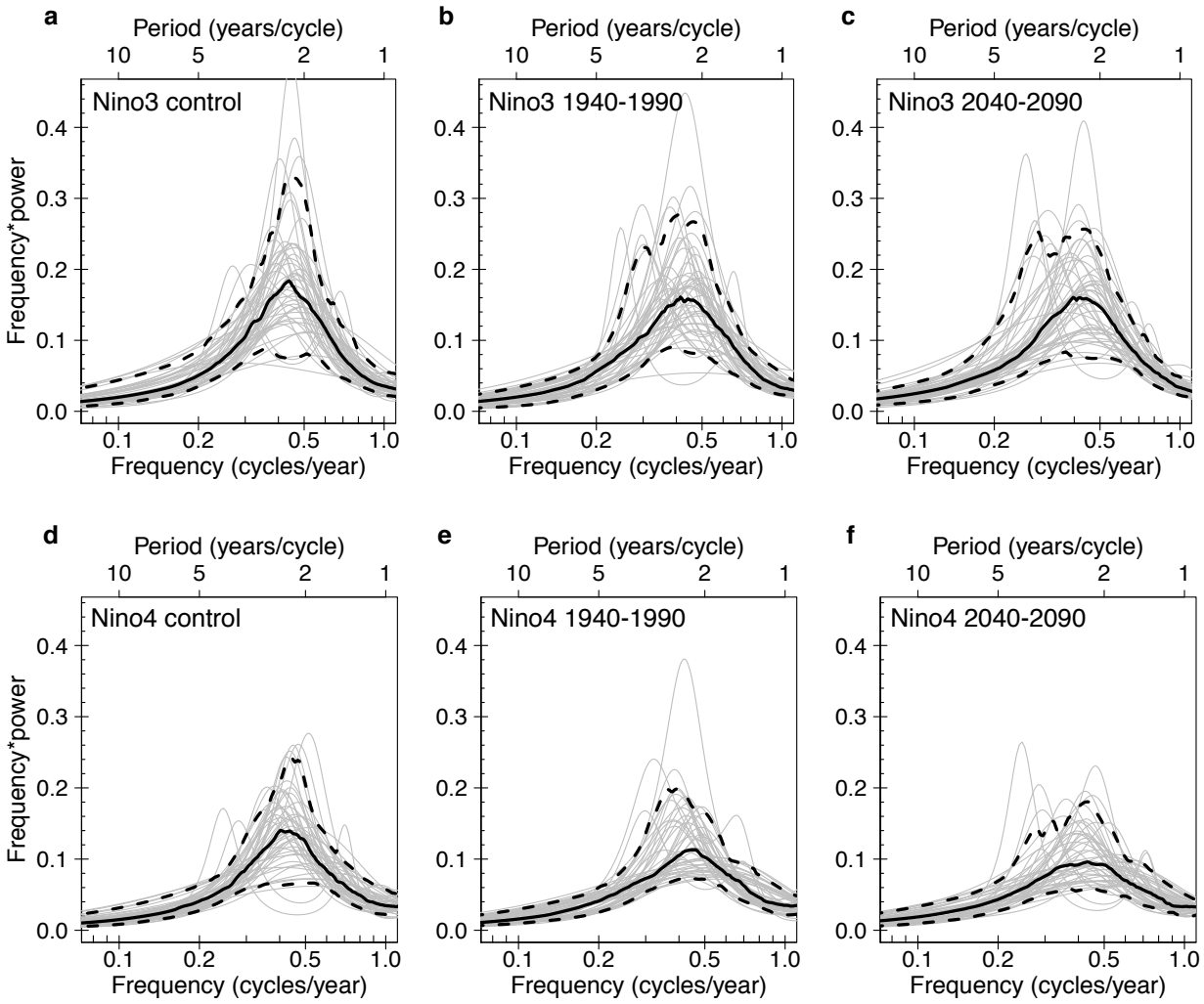

Figure 10 | Maximum entropy power spectra for different ENSO indices and forcing regimes. **a**, **b**, **c**, Trend-removed Niño3 index in different forcing regimes; **d**, **e**, **f**, Trend-removed Niño4 index in different forcing regimes. **a**, **d**, 50 50-year sections of the unforced, control CESM simulation; **b**, **e**, the 50 CESM ensemble members (1940-1990); **c**, **f**, the 50 CESM ensemble members (2040-2090). Individual members are shown in grey, dashed curves are the 5th and 95th percentiles, and the solid black curve is the median.

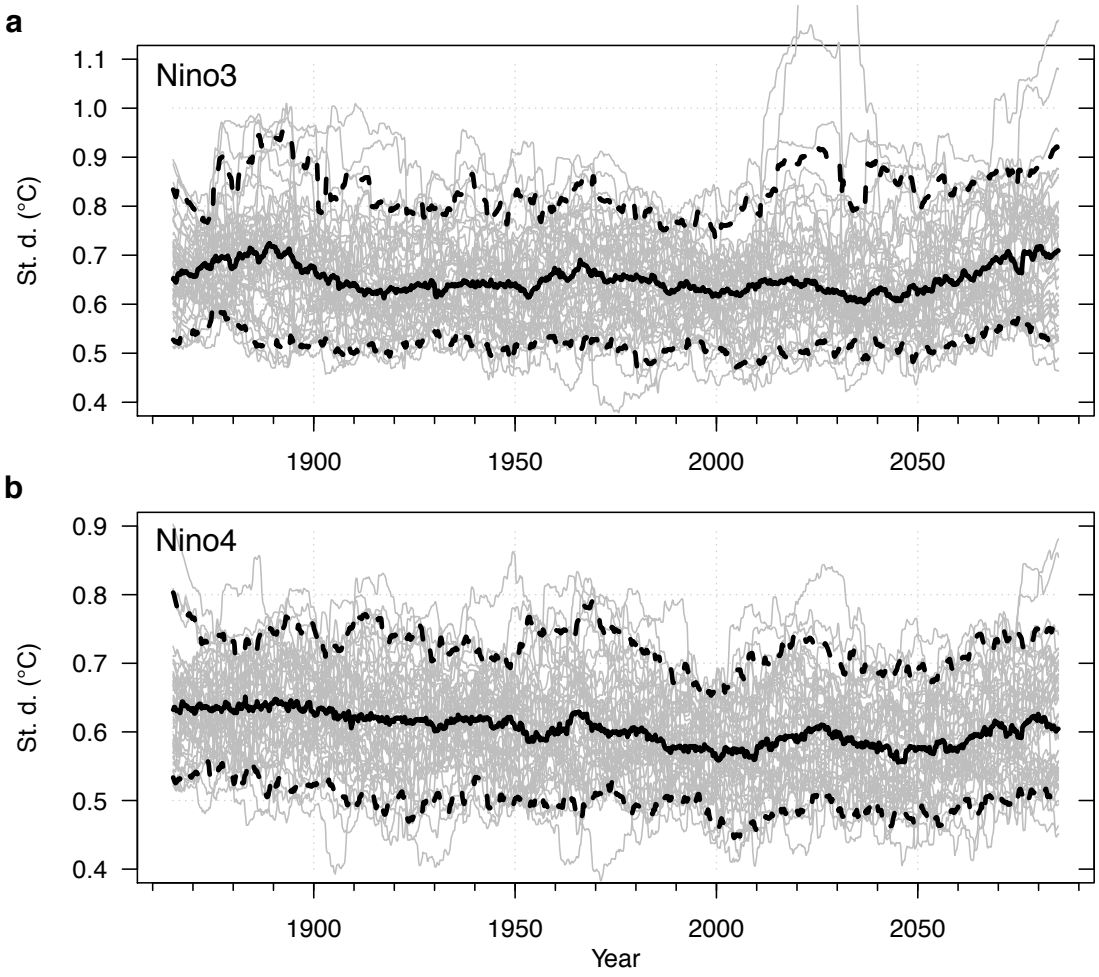

Figure 11 | 20-year running standard deviation for different ENSO indices. a, Trend-removed Niño3 index; **b**, trend-removed Niño4 index. The 50 CESM ensemble members are shown in grey, dashed lines are the 5th and 95th percentiles, and the solid black line is the median.

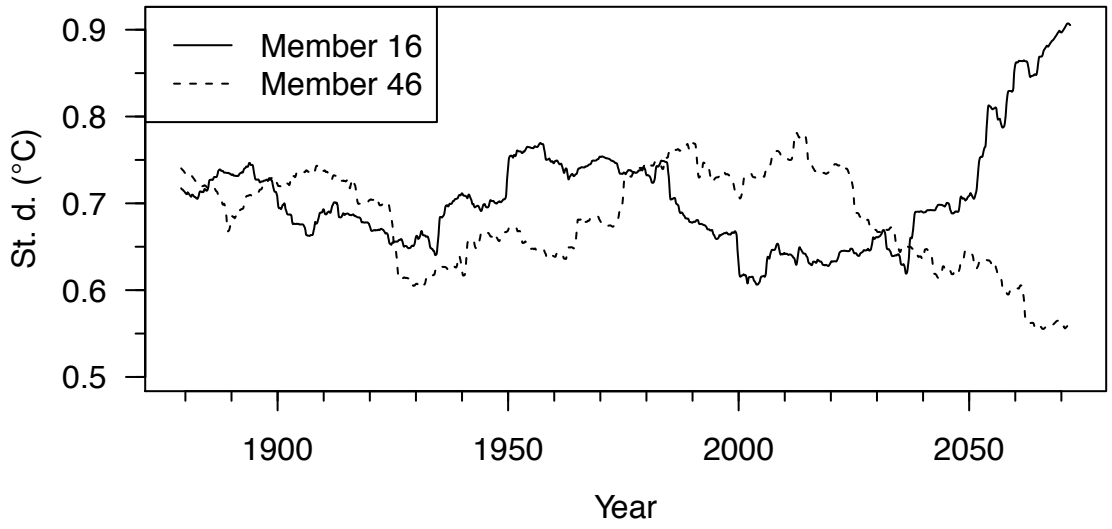

Figure 12 | 50-year running standard deviation of the trend-removed Niño 3.4 index for two sample ensemble members. These members are chosen to highlight the different 21st century trends possible under the same forcing.

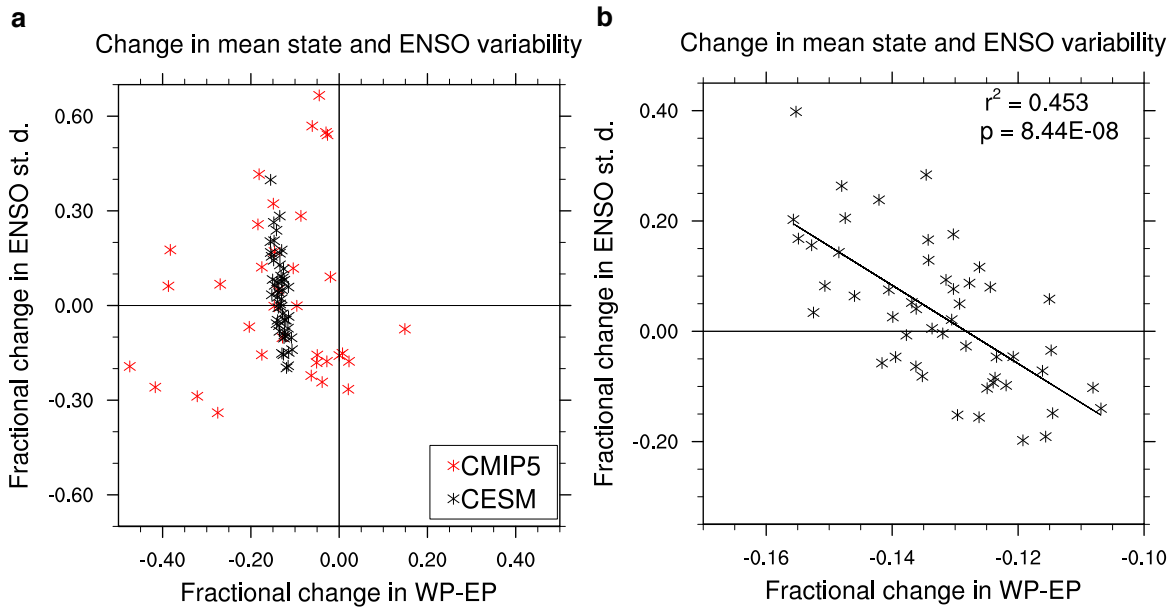

Figure 13 | Changes in the equatorial zonal temperature gradient and Niño3.4 standard deviation. a, Fractional change in Western equatorial Pacific (5°S-5°N, 120°E-170°E) minus Eastern equatorial Pacific (5°S-5°N, 150°W-90°W) SST vs. fractional change in trend-removed Niño3.4 standard deviation from 1940-1990 to 2040-2090. In red are the 35 CMIP5 ensemble members; in black are the 50 CESM ensemble members. Negative fractional changes in WP-EP indicate preferential warming in the east. **b**, Same as **a**, but only showing the CESM ensemble members. The listed p-value is that of the null hypothesis of zero slope.

TABLES

	Control mean	2040-2090 mean	|Control-2040-2090|	p-value	95% Detection sensitivity	95% Detection sensitivity (%)
St. dev. (°C)	0.701	0.698	0.003	0.823	0.026	3.7
Skewness	0.263	0.325	0.058	0.308	0.12	45.9
El Niño count	11.14	11.28	0.14	0.732	0.81	7.3
La Niña count	12.54	12.82	0.28	0.516	0.85	6.8

Table 1 | Welch's two-sample t-test results comparing ENSO properties of the control CESM ensemble and the 2040-2090 ensemble. The p-value gives the two-sided probability of the given difference in means under the null hypothesis of equal means. A p-value less than 0.05 would mean rejection of the null hypothesis with 95% confidence. The 95% detection sensitivity is the approximate difference in means that would be needed to reject the null hypothesis with 95% confidence. The detection sensitivity is calculated assuming 98 degrees of freedom and that the standard deviation of the given property is constant. The fractional detection sensitivity of skewness is much worse than that of the other properties because the fractional distributions of skewness are much larger (Fig. 7).